

Developmental appearance of factors that bind specifically to *cis*-regulatory sequences of a gene expressed in the sea urchin embryo

Frank J. Calzone, Nadine Thézé, Pierre Thiebaud,¹ Ronald L. Hill, Roy J. Britten, and Eric H. Davidson

Division of Biology, California Institute of Technology, Pasadena, California 91125 USA

Previous gene-transfer experiments have identified a 2500-nucleotide 5' domain of the *CyIIIa* cytoskeletal actin gene, which contains *cis*-regulatory sequences that are necessary and sufficient for spatial and temporal control of *CyIIIa* gene expression during embryogenesis. This gene is activated in late cleavage, exclusively in aboral ectoderm cell lineages. In this study, we focus on interactions demonstrated *in vitro* between sequences of the regulatory domain and proteins present in crude extracts derived from sea urchin embryo nuclei and from unfertilized eggs. Quantitative gel-shift measurements are utilized to estimate minimum numbers of factor molecules per embryo at 24 hr postfertilization, when the *CyIIIa* gene is active, at 7 hr, when it is still silent, and in the unfertilized egg. We also estimate the binding affinity preferences (K_d) of the various factors for their respective sites, relative to their affinity for synthetic DNA competitors. At least 14 different specific interactions occur within the regulatory regions, some of which produce multiple DNA-protein complexes. Values of K_d range from $\sim 2 \times 10^4$ to $\sim 2 \times 10^6$ for these factors under the conditions applied. With one exception, the minimum factor prevalences that we measured in the 400-cell 24-hr embryo nuclear extracts fell within the range of 2×10^5 to 2×10^6 molecules per embryo, i.e., a few hundred to a few thousand molecules per nucleus. Three developmental patterns were observed with respect to factor prevalence: Factors reacting at one site were found in unfertilized egg cytoplasm at about the same level per egg or embryo as in 24-hr embryo nuclei; factors reacting with five other regions of the regulatory domain are not detectable in egg cytoplasm but in 7-hr mid-cleavage-stage embryo, nuclei are already at or close to their concentrations in the 24-hr embryo nuclei; and factors reacting with five additional regions are not detectable in egg cytoplasm and are low in 7-hr embryo nuclei, i.e., $\leq 10\%$ per embryo of the level they attain in 24-hr embryo nuclei. The rise in concentration of factors of the latter class could provide the proximal cause for the temporal activation of the *CyIIIa* gene at the early blastula stage.

[Key Words: Transcription factors; development; sea urchin]

Received April 4, 1988; revised version accepted July 6, 1988.

The *CyIIIa* cytoskeletal actin gene of *Strongylocentrotus purpuratus* is transcriptionally activated at the early blastula stage of embryonic development (Shott et al. 1984; Lee 1986; Lee et al. 1986; Hickey et al. 1987). From the earliest point at which accumulation of its transcripts becomes detectable and throughout embryogenesis, the *CyIIIa* gene is expressed exclusively in the cell lineages that give rise to the aboral ectoderm of the advanced embryo and larva (Angerer and Davidson 1984; Cox et al. 1986). This structure is formed from six clonal lineages, the progenitors of which segregate from other founder cells at the third to sixth cleavages (Cameron et al. 1987; R. Cameron, unpubl.). *CyIIIa* is one of a number

of diverse genes that are now known to be activated in a lineage-specific manner toward the end of cleavage, which appears to be an important transition point in sea urchin development (for review, see Davidson 1986, pp. 238–246). The precise spatial and temporal pattern of *CyIIIa* gene expression has engendered a series of studies directed at the molecular basis of its regulation. Both maternal and embryonic positional information are evidently utilized in this regulatory process. Thus, the orientation of the oral–aboral axis of the embryo is specified epigenetically after fertilization but parallel to the horizontal plane of third cleavage, which is always set perpendicularly to the preformed animal–vegetal axis of the egg (for review, see Davidson 1986, pp. 494–504). Though the polarity of the oral–aboral axis has been established by third cleavage (Czihak 1963; Davidson

¹ Present address: Laboratoire de biochimie du développement, Faculté des sciences, Beaulieu, 35040 Rennes, France.

1986; Cameron et al. 1987), expression of the set of genes expressed specifically in the aboral ectoderm, including *CyIIIa*, does not begin until five to six further cycles have elapsed (Bruskin et al. 1981, 1982; Hickey et al. 1987; for review, see Davidson 1986). Thus, the localization of *CyIIIa* activity reflects an epigenetic decision made very early in development, which is ultimately transformed into a differential pattern of gene expression.

Gene transfer and in vivo competition experiments have identified a 2.5-kb regulatory region upstream of the *CyIIIa* cap site that is sufficient for correct developmental expression. When covalently associated with the chloramphenicol acetyltransferase (CAT) reporter gene, this region directs temporal (Flytzanis et al. 1987) and spatial (Hough-Evans et al. 1987, 1988; Franks et al. 1988) expression of CAT mRNA on the same schedule and in the same cell types as for the endogenous *CyIIIa* genes. Previous results suggest that the *CyIIIa* regulatory region includes a number of independent sites where functionally significant *cis-trans* interactions take place. Thus, (1) various deletions in the 5'-regulatory region of the *CyIIIa*·CAT fusion decrease the activity of the construct to different extents in gene transfer experiments (Flytzanis et al. 1987). (2) The activity of the *CyIIIa*·CAT fusion construct can be saturated by introduction of a sufficient number of copies, and coinjection of excess molecules of the regulatory region per se depresses activity almost stoichiometrically, evidently by competing with *cis* sequences linked to the CAT reporter gene for factors present in the embryonic nuclei in limited availability (Flytzanis et al. 1987; Livant et al. 1988). (3) By coinjection of molar excesses of given subfragments of the 2-kb regulatory region, together with the complete *CyIIIa*·CAT fusion construct, several nonoverlapping sequences have been identified with which interactions are apparently required for normal expression (R. Franks, unpubl.) (4) On injection of the *CyIIIa*·CAT fusion into eggs of a distant sea urchin species, correct temporal regulation is preserved, whereas regulation of the spatial expression pattern fails. This last observation suggests functional separability of interactions controlling temporal and spatial ex-

pression. To characterize the DNA-protein interactions that occur in the regulatory region and to determine their individual specificities under equilibrium conditions, we have carried out the in vitro binding studies described in this paper. We find at least 14 individual sites within the regulatory domain where specific DNA-protein interactions occur. The exact locations of these sites and an analysis of the sequences involved are to be presented elsewhere. Our focus in this study is developmental. The in vitro reactions provide minimum estimates of the concentrations in crude extracts of the DNA-binding factors that react with the various subregions of the *CyIIIa* regulatory domain and, thus, of changes in their prevalence during development. Measurements were carried out on nuclear extracts of 24-hr embryos, in which the *CyIIIa* gene is functioning actively (Shott et al. 1984; Cox et al. 1986; Lee 1986; Lee et al. 1986; Hickey et al. 1987); on nuclear extracts of 7-hr mid-cleavage embryos, which are still one or two division cycles prior to activation of the *CyIIIa* gene; and on cytoplasmic extracts of unfertilized eggs. Several distinct developmental patterns in binding factor concentration emerge that are relevant to the possible functional values of the individual *cis*-sequence elements at which these factors react.

Results

The *CyIIIa* regulatory region

A map of the *CyIIIa* genomic sequence extending 2300 nucleotides upstream from the transcriptional initiation site is shown in Figure 1, where we present the nomenclature utilized in this study. Restriction enzyme subfragments that in gel shift assays display specific interactions with nuclear protein extracts from 24-hr embryos, are labeled P1–P8, our designations for the respective sets of DNA-binding factors. Those subfragments that lack any such designation in Figure 1 have been shown specifically not to contain any sequence elements that react with these nuclear extracts under the conditions of specificity that we imposed. The set of subcloned probes with which the gel-shift experiments

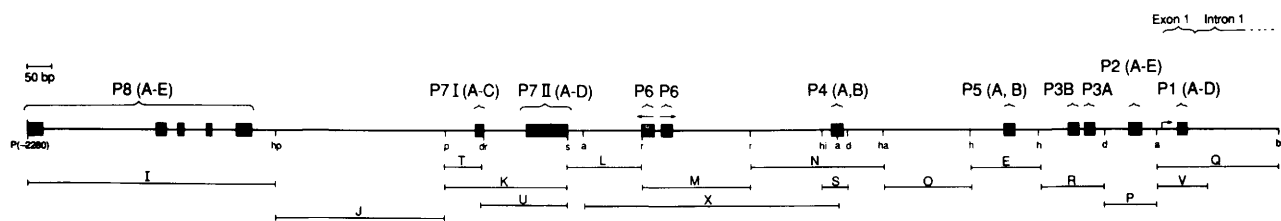


Figure 1. Map of *CyIIIa* *cis*-acting control elements. Black boxes indicate the location of the sequence elements in the 5'-flanking region of the *CyIIIa* gene that bind embryo nuclear or egg cytoplasmic proteins. The map summarizes results of extensive deletion and oligonucleotide competition experiments, using gel retardation assays to identify the probe binding site in the complexes shown in Results (N. Thézé et al., unpubl.). In some cases, sequences required for binding did not completely overlap regions protected from DNase I in footprint studies. The complexes named in Results for each binding site are indicated above the map, as are the locations of transcription initiation and the exon 1/intron 1 boundary. There are two adjacent copies of the P6-binding site, located in inverse orientation. The probes referred to in Results are indicated below the gene map. Restriction sites used to construct the subcloned probes designated below the map are: *Ava*I (a), *Bam*HI (b), *Dde*I (d), *Dra*I (dr), *Hind*III (h), *Hae*III (ha), *Hinf*I (hi), *Hpa*II (hp), *Pst*I (p), *Rsa*I (r), and *Sau*3A (s).

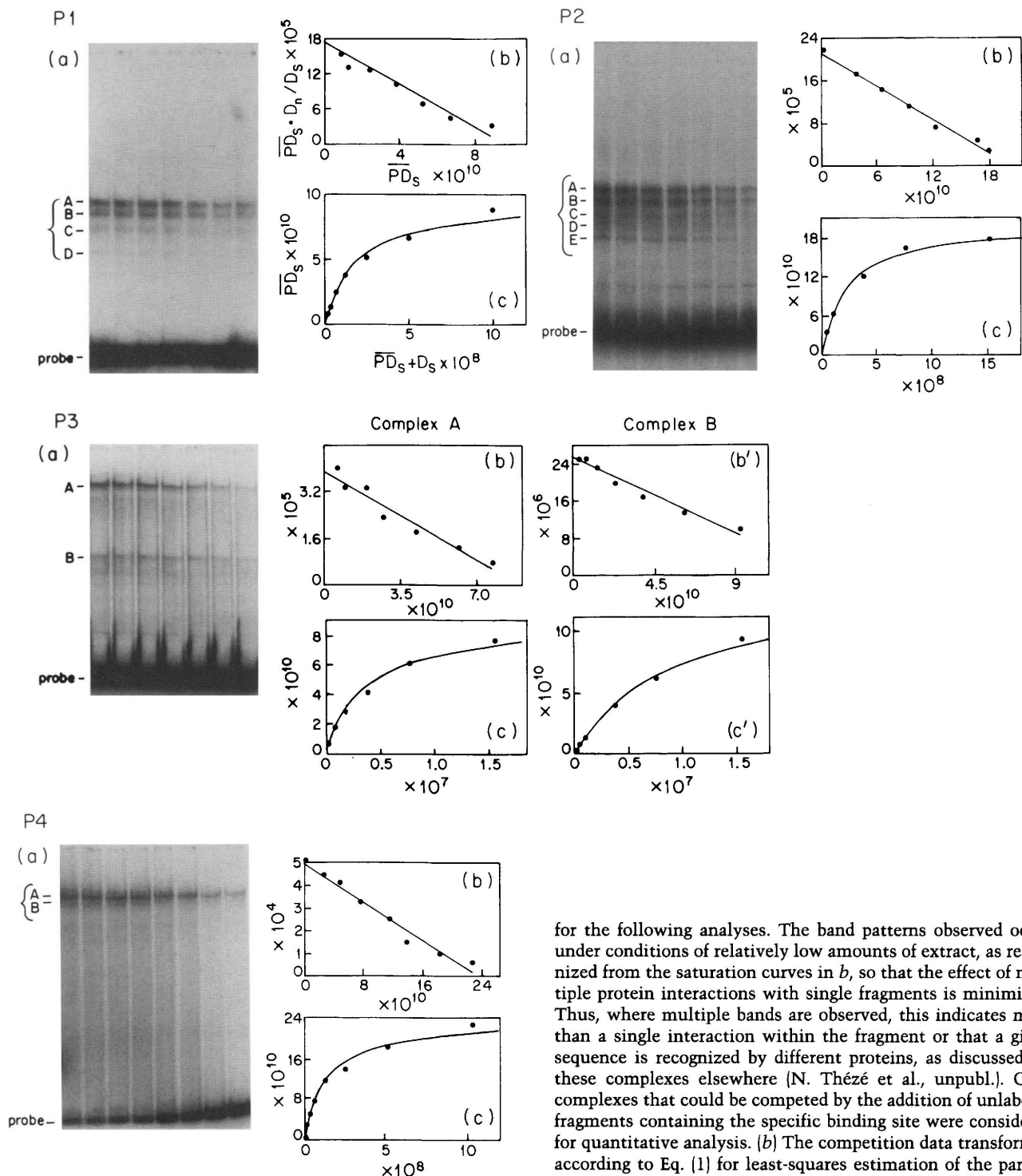
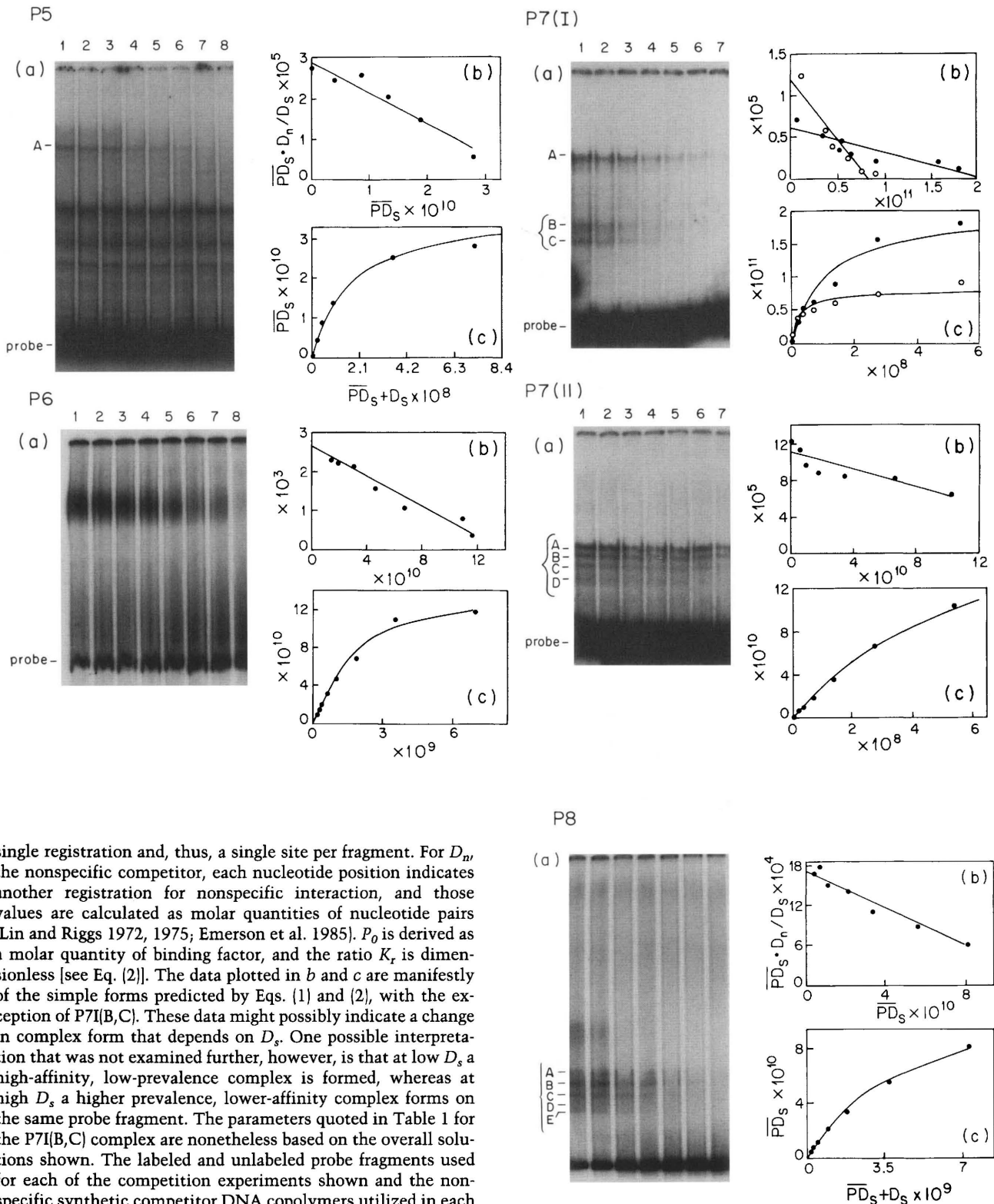


Figure 2. Gel-shift competitions for 24-hr embryo nuclear DNA-binding factors and measurements of K_r and P_0 . Nuclear extracts were reacted with the various probes (see Fig. 1; Table 1), together with increasing concentrations of unlabeled specific competitor DNA, as described in text. Autoradiographs of each gel-shift competition series are shown in *a*. The location of the free probe and the probe complex(es) is indicated in each panel. Several forms of gel-shift complex were reproducibly observed in many cases. Brackets indicate those forms summed

for the following analyses. The band patterns observed occur under conditions of relatively low amounts of extract, as recognized from the saturation curves in *b*, so that the effect of multiple protein interactions with single fragments is minimized. Thus, where multiple bands are observed, this indicates more than a single interaction within the fragment or that a given sequence is recognized by different proteins, as discussed for these complexes elsewhere (N. Thézé et al., unpubl.). Only complexes that could be competed by the addition of unlabeled fragments containing the specific binding site were considered for quantitative analysis. (b) The competition data transformed according to Eq. (1) for least-squares estimation of the parameters P_0 and K_r (see text). Experimental data is indicated by circles here and in *c*, and the line is the best fit least-squares solution, assuming the indicated functions. (c) The relation between quantity of specific complex formed (PD_s) and quantity of total specific binding site present ($PD_s + D_s$), for each of the binding factors (P1–P8) indicated at left. The values of PD_s used to generate the curves shown (by a spline interpolation routine) were calculated according to Eq. (2), using the parameters extracted from the analyses shown in the respective *b* panels. D_s and PD_s are expressed as molar quantities of sites, assuming a



single registration and, thus, a single site per fragment. For D_n , the nonspecific competitor, each nucleotide position indicates another registration for nonspecific interaction, and those values are calculated as molar quantities of nucleotide pairs (Lin and Riggs 1972, 1975; Emerson et al. 1985). P_0 is derived as a molar quantity of binding factor, and the ratio K_r is dimensionless [see Eq. (2)]. The data plotted in *b* and *c* are manifestly of the simple forms predicted by Eqs. (1) and (2), with the exception of P7I(B,C). These data might possibly indicate a change in complex form that depends on D_s . One possible interpretation that was not examined further, however, is that at low D_s a high-affinity, low-prevalence complex is formed, whereas at high D_s a higher prevalence, lower-affinity complex forms on the same probe fragment. The parameters quoted in Table 1 for the P7I(B,C) complex are nonetheless based on the overall solutions shown. The labeled and unlabeled probe fragments used for each of the competition experiments shown and the nonspecific synthetic competitor DNA copolymers utilized in each case are listed in Table 1. Here, the first entry for each binding factor is the experiment shown. Nonspecific competitors were present in the experiments shown at $\geq 4 \times 10^4$ the mass of the labeled probe DNAs, as indicated in Materials and methods. For P7I closed circles indicate complex A and open circles complexes B and C (a; ●) Complex A; (○) complexes B and C.

were carried out is shown beneath the map in Figure 1. Transcription of the CyIIIa gene begins within the region included in Probe V which overlaps the first exon of the gene. The initiation site of this gene was identified by primer extension in an earlier study (Akhurst et al. 1987). The first exon is 75 nucleotides in length and contains only nontranslated leader sequences. The gene is then interrupted by a 2.2-kb leader intron. The possibility that regulatory sites exist within this intron has not been explored, as the standard CyIIIa · CAT fusion construct with which the gene transfer studies cited in the introductory section were carried out all include this intron. Because deletions that remove sequences upstream of the initiation site abolish activity, we can conclude only that if any regulatory interactions occur with sequence elements within the leader intron, they are not in themselves sufficient. Nor is there any detectable competitive effect when a large molar excess of a fragment including most of the leader intron sequence is coinjected with the CyIIIa · CAT fusion, as there is when molar excess of fragments containing the upstream regulatory domain are introduced (Livant et al. 1988; R. Franks, unpubl.).

Minimum prevalences and specific site affinities for DNA-binding factors of 24-hr embryo nuclei

In this section we derive from gel shift data estimates of P_0' , the number of nuclear molecules of each DNA-binding factor per embryo, and of K_r , which measures the preference each factor displays for its CyIIIa-binding site, compared with its affinity for synthetic copolymer DNAs ('nonspecific competitor' DNA). K_r is defined as the ratio of the equilibrium constant for the complex between the factor and its specific site in the CyIIIa regulatory region to that for the complex between the factor and the nonspecific competitor DNA (i.e., $K_r = K_s/K_n$). The method we have used to extract these parameters is derived from a treatment of Emerson et al. (1985) obtained from elementary equilibrium assumptions, given the condition that the concentrations of specific site (D_s), of nonspecific competitor DNA sites (D_n), and of molecules of a given factor in the system (P_0), are such that there will be essentially no free factor protein. That is, because of the proclivity of sequence specific DNA-binding proteins to associate with any DNA, though forming much lower stability complexes relative to those generated at the specific sites, essentially all factor molecules are assumed to be bound either to the specific probe DNA or the nonspecific competitor DNA present in the reactions. The molar concentration of specific complex formed at equilibrium (PD_s) is related to the other parameters defined as in Eq. (1):

$$\frac{PD_s \cdot D_n}{D_s} = K_r(P_0 - PD_s) \quad (1)$$

This provides a linear function conveniently fit by least-squares procedures, in which the ordinate intercept is P_0K_r , and the slope is $-K_r$ (Emerson et al. 1985). The saturation function describing the amount of specific

complex present at equilibrium as the amount of specific site increased is then

$$PD_s = \frac{P_0K_rD_s}{D_n + K_rD_s} \quad (2)$$

To obtain values of P_0 and K_r for the CyIIIa-binding factors, we carried out gel-shift experiments (Fried and Crothers 1981) with a fixed amount of labeled specific DNA probe and unlabeled competitor DNA, and increased D_s by addition of unlabeled specific DNA probe in successive samples. D_s in the following analysis is the molar concentration of unoccupied specific sites available in the reaction. Over an appropriate range of PD_s/P_0 values, increase of unlabeled probe will result in competition and a decrease in formation of complex with the homologous labeled probe. Autoradiographs displaying a gel-shift competition series for each of the binding factors are reproduced in Figure 2a. Note that multiple complexes were formed with a single probe fragment in several cases [e.g., P1(A–D) in Fig. 2; see legend]. The quantities of total (i.e., labeled plus unlabeled) specific complex were calculated from the distribution of radioactivity between the free and bound probe forms. These data were then utilized to provide estimates of P_0 and K_r , according to Eq. (1), as illustrated in Figure 2b. The derived P_0 and K_r values could then be utilized to generate the saturation functions shown in Figure 2c, by application of Eq. (2). In a number of cases, bands were observed in the gel-shift experiments that failed to diminish in radioactivity as unlabeled probe was added. These complexes that cannot be competed are of low specific affinity and are formed by proteins of very high prevalence relative to those responsible for the specific complexes. Nonspecific complexes were ignored in the following study, though we note that at least some of them occur only at particular locations within the probe sequences (data not shown).

Results from these and many additional experiments, obtained with several independent nuclear extracts, are listed in Table 1, where the prevalence of each factor is given per embryo (P_0'). In these analyses where multiple bands appear, we consider the protein or proteins reacting with a given probe as 'a factor' if these bands display similar competition profiles. The P_0' values were calculated from the molar prevalences, (P_0), of the respective factors in the reaction mixtures, derived from least-squares analyses such as that shown in Figure 2b (see footnote to Table 1). The P_0' values are minimal because they do not include recovery efficiencies, which are unknown. In Table 1 it can be seen that the multiple determinations for each factor display a satisfactory consistency in respect to both P_0' and K_r , even where the measurements were carried out on independent extracts. Furthermore, no systematic correlation exists between K_r and P_0' values. We therefore regard the differences among the factors in respect to both their prevalences and their specific affinities as real and characteristic distinctions. Thus, e.g., the prevalences of P2 and P6 are at least 1×10^6 to 2×10^6 molecules per embryo, whereas the two P7(I) factors are 20–90× less abundant. The other factors are all present at intermediate levels of at

Table 1. P_0 and K_r measurements for 24-hr nuclear extracts

Factor	P_0^a (molecules/embryo)	K_r^a	Nonspecific competitor ^b	Extract ^c	Probe ^d	Specific competitor ^d
P1(A-D)	4.4×10^5	3.0×10^5	poly[d(AT)]	4	V	Q
	4.7×10^5	1.9×10^5	poly[d(AT)]	1	V	Q
P2(A-E)	1.9×10^6	1.1×10^5	poly[d(AT)]	1	P	P
	2.7×10^6	9.6×10^4	poly[d(AT)]	2	P	P
P3(A)	2.4×10^5	3.1×10^4	poly[d(AT)]	1	R	R
	4.8×10^5	4.4×10^4	poly[d(AT)]	1	R	R
P3(B)	7.5×10^5	1.7×10^4	poly[d(AT)]	1	R	R
P4(A,B)	2.1×10^5	2.8×10^5	poly[d(AT)]	5	S	N
	2.3×10^5	2.7×10^5	poly[d(AT)]	5	S	N
P5(A)	2.1×10^5	7.8×10^4	polyd(I)/polyd(C)	1	E	E
	8.2×10^4	6.7×10^4	polyd(I)/polyd(C)	1	E	E
	2.2×10^5	7.6×10^4	polyd(I)/polyd(C)	4	E	E
P6	1.2×10^6	1.9×10^6	poly[d(AT)]	1	X	X
	1.2×10^6	1.4×10^6	poly[d(AT)]	1	X	M
	1.5×10^6	1.5×10^6	poly[d(IC)]	1	X	X
	1.8×10^6	7.1×10^5	poly[d(I)/polyd(C)]	1	X	X
	1.2×10^6	1.4×10^6	poly[d(AT)]	3	X	X
P7I(A)	6.2×10^4	3.3×10^5	poly[d(AT)]	4	T	K
P7I(B,C)	2.5×10^4	1.6×10^6	poly[d(AT)]	4	T	K
P7II(A-D)	7.4×10^5	4.7×10^4	poly[d(AT)]	4	U	K
P8(A-C)	8.0×10^5	9.1×10^5	poly[d(AT)]	1	I	I
	4.4×10^5	1.1×10^6	poly[d(AT)]	1	I	I
	6.3×10^5	9.7×10^5	poly[d(AT)]	4	I	I

^a P_0' and K_r values were extracted from gel-shift competition data such as that shown in Fig. 2. The initial entry for each binding factor is the particular experiment illustrated for that factor in Fig. 2. These least-squares analyses directly provide values of the parameters P_0 , the molar concentration of the binding factor in the in vitro reaction mixture. P_0' , the number of molecules of factor per embryo (i.e., in the nuclear compartment), listed here, is obtained from P_0 by applying the volume of the gel-shift reaction mixtures, typically 10 μ l, and the number of embryos represented by the amount of nuclear protein extract added, typically 6×10^3 to 30×10^3 embryos (see Materials and methods). For example, for the experiment shown in Fig. 2 for P6, 0.3 μ l of extract, representing 1.8×10^4 embryos, was added to 0.5 ng of probe, plus 6 μ g of poly[d(AT)] competitor, in a total volume of 10 μ l. The values of P_0' shown are minimal ones, as we cannot estimate the efficiency with which the factors were recovered from the embryo nuclei.

^b Different nonspecific DNA competitors were utilized as indicated. In some cases, the identity of the competitor made little difference, but in others, e.g., P5, the competitor shown worked well, whereas the alternative polymers yielded weak bands or otherwise failed to provide clear results.

^c Extracts 1–5 were prepared independently from different batches of embryos.

^d See Fig. 1 for map and identities of the probe and competitor fragments indicated.

least 10^5 – 10^6 per embryo. We do not yet know which of these factors, if any, are confined to the aboral ectoderm cell progenitors of the 24-hr embryo or to nonaboral ectoderm cell types. About 40% of the cells of the 24-hr embryo are aboral ectoderm precursors, and there is thus an uncertainty of a factor of 2.5 in the actual minimum concentrations of these factors per nucleus. If present in all nuclei of the 400-cell embryo, there would be at least 5000 molecules per nucleus of a factor such as P2 (2×10^6 per 10 embryo) and at least 250 molecules per nucleus of a factor such as P5 (1×10^5 per embryo).

The specific affinities (K_r) derived for the various reactions also differ by large factors. The least specific are the complexes formed by P3(A), P3(B), and P7(II), which, however, still prefer their *CyIIa*-binding sites to the nonspecific competitor DNA by ratios of 2×10^4 to 5×10^4 . The other factors display K_r values in the range 10^5 – 10^6 . It is interesting to note that in two of the three cases studied earlier by an exonuclease competition method, the K_r values obtained with natural satellite

DNA as the nonspecific competitor were indistinguishable from those obtained against the synthetic copolymeric DNAs listed in Table 1. Thus, Calzone et al. (1987) reported a K_r value for P3(A) of 4.6×10^4 and a K_r value for P2 of 1.1×10^5 (cf. Table 1). However, such is not always expected to be the case, because some nonspecific competitors may inadvertently bear resemblance to the specific site for a given factor as, e.g., poly[d(AT)] does for the sites bound by *lac* repressor (Lin and Riggs 1972). Thus, as noted in the footnote to Table 1, some determinations displayed sensitivity to the nature of the synthetic DNAs used as nonspecific competitor in our experiments. The K_r reported by Calzone et al. (1987) for P1 is 10–20 \times lower than that listed in Table 1. However, the significance of this is unclear, as there is indirect evidence for interactions at several nearby locations that affect K_r for P1 (F. Calzone, unpubl.), and the probe used in earlier and present experiments was not the same. In any case, it is clear from the K_r values listed in Table 1 that the binding-factor-site

interactions observed are highly sequence specific. Discounting accidental homologies, the nonspecific competitor DNA controls should, at minimum, correct for binding interactions that require only the sequence-independent aspects of native DNA structure.

Stability of *CyIIIa* DNA-protein complexes

The measurements shown in Figure 3 provide direct estimates of the disassociation rate constants (k_d), under the in vitro reaction conditions utilized, for two of the DNA-protein complexes. We chose representative examples, i.e., P4(AB), for which the measured K_r is $\sim 2.8 \times 10^5$, and P6, for which it is $\sim 1.2 \times 10^6$ (Table 1). The nuclear extracts were incubated for 10 min with the respective probes for these factors (see Table 1) as above. Aliquots of the reaction mixture were then incubated with a large excess of unlabeled competitor DNA and were loaded on a running gel after the intervals shown on the abscissa. Once having entered the gel, effective complex dissociation is arrested or greatly retarded (Fried and Crothers 1981). Figure 3 indicates a half-life of ~ 14.4 min for the P4 complex and ~ 47.8 min. for the P6 complex. Thus, k_d for P4 is 8.0×10^{-4} /sec and k_d for P6 is 2.4×10^{-4} /sec (i.e., $k_d = \ln 2/t_{1/2}$). The ratio of these values is ~ 3.3 , and were the stability of the specific complexes the sole determinants of their equilibrium behavior, a ratio of 4.3 would have been expected from the average K_r values. Barring coincidence, this result would suggest that the association rate (k_a) for these two factors must be approximately the same and so must the equilibrium dissociation constants of their as-

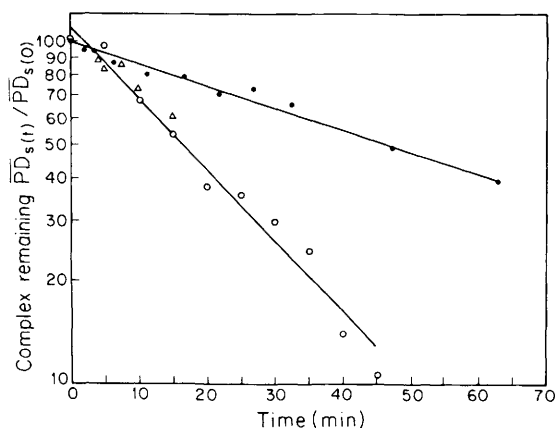


Figure 3. Stability determination for specific complexes formed by P4 and P6 in crude extracts. Measurements were carried out at 0°C. Reactions with probes S (P4) and M (P6) were set up as for the experiments presented in Table 1 and Figure 2. The concentrations of labeled probe in each reaction were 1.1×10^{-9} M for P4 ($PD_s/P_0 \sim 0.05$) and 2.12×10^{-10} M for P6 ($PD_s/P_0 \sim 0.44$). After a 10-min incubation, aliquots of the reaction were transferred to a tube containing a 200- to 400-fold excess of unlabeled probe and incubated for the times indicated. The fraction of complex remaining was determined by gel electrophoresis, as described for the experiments shown in Fig. 2. (○, △) P4; (●, ▲) P6.

sociation with the nonspecific DNA competitor [i.e., $K_r(P6)/K_r(P4) = K_s(P6)/K_s(P4)$ if $K_n(P4) = K_n(P6)$]; and because $K_s = k_a/k_d$, $K_s(P6)/K_s(P4) = k_d(P4)/k_d(P6)$ if $k_a(P4) = k_a(P6)$.

To obtain estimates of k_a and K_n , i.e., in the presence of the crude nuclear extracts, we also measured K_{Ds} , i.e., K_s^{-1} , the equilibrium dissociation constant, for the P4 reaction. Baker et al. (1986) derived an 'apparent equilibrium constant, K_{app} ', for conditions where D_n is not necessarily in excess, which is given by Eqs. (3) and (4):

$$\frac{PD_s}{D_s} = K_{app}(P_0 - PD_s) \quad (3)$$

$$\frac{1}{K_{app}} = \frac{1}{K_s} + \frac{1}{K_r} D_n \quad (4)$$

In practice, given the value of P_0 (and K_r) obtained from D_s titrations as in Figure 2, a D_n titration is carried out, and K_{app}^{-1} calculated as $(P_0 - PD_s)/(PD_s/D_s)$ (Eq. 3). This yields estimates of K_{Ds} (i.e., the ordinate intercept), and K_n from the slope ($K_n/K_s = K_r^{-1}$). In our experience, a source of concern with this procedure is that the molar concentration of nonspecific competitor sites required may be sufficiently high under the conditions of these assays that the viscosity is increased and, thus, the reactions must be diluted. We found that the relative affinity of the factor for the specific vs. nonspecific DNA for P6 is unusually high, and the value of K_r recovered from the D_n titration was 2–3 \times greater than the K_r values listed in Table 1 for P6. K_{Ds} for P6 is 1.02×10^{-9} M (data not shown; average of two measurements with correlation coefficients of 0.95 and 0.98). For P4 the value of the equilibrium dissociation constant K_{Ds} for the specific complex according to Eq. (4), is $K_{Ds} = 5.8 \times 10^{-9}$ M (data not shown; the correlation coefficient for this analysis was 0.99). The value of K_r derived from this experiment was 2.7×10^5 , i.e., exactly the same as listed in Table 1. The value of k_a observed in the crude extracts (0°C) would then be 1.4×10^5 M⁻¹/sec (i.e., $k_a = k_d/K_{Ds}$). From the D_n titration we obtain for the reaction of P4 with the poly[d(AT)] competitor, an equilibrium dissociation constant $K_{Dn} = 1.8 \times 10^{-3}$ M, and, if we assume the same association rate constant for the specific and nonspecific reactions (Riggs et al. 1970; Jobe et al. 1974), k_d for the nonspecific reaction would be 2.5×10^2 /sec; i.e., the half-life for the nonspecific DNA complex must be only about 3 msec.

Maternal factors that interact with the *CyIIIa* regulatory region

To determine whether any of the DNA-binding factors would be of maternal rather than zygotic origin, soluble proteins of unfertilized eggs were fractionated on a heparin-agarose column (for details, see Materials and methods), and proteins that bound to this column were tested for specific interactions with the regulatory region probes. Recovery efficiency for the heparin-agarose procedure is in our hands 50–80% (unpublished mea-

surements). The egg protein extracts were obtained from centrifugal supernatants of egg homogenates (for details, see Materials and methods), and thus any DNA-binding factors stably associated with the pronuclei would not have been included. Any germinal vesicle proteins released at breakdown and retained in the cytoplasm might have been recovered, however (germinal vesicle breakdown occurs spontaneously at the end of the growth phase of oogenesis in the sea urchin, and the matured eggs are then retained in the ovary for up to several weeks; for review, see Davidson, 1986). The egg extracts were reacted with the regulatory region probes shown in Figure 1, and only in those subfragments where interactions occur with the 24-hr nuclear factors P3, P4, and P5

did specific, high-affinity complexes with any maternal factors form. The gel-shift patterns obtained with these maternal proteins are shown in Figure 4, where they can be compared with the patterns obtained with the nuclear extracts from 7- and 24-hr embryos. The most interesting qualitative observation illustrated in Figure 4 is that with one minor exception, the gel-shift complex patterns generated by the maternal proteins are all different from those observed in 7- and 24-hr embryo nuclear extracts, whereas the sets of P3, P4, and P5 complexes formed by 7- and 24-hr nuclear extracts are largely similar.

Although the maternal P3 and P4 factors are present at too low a concentration to permit quantitative measure-

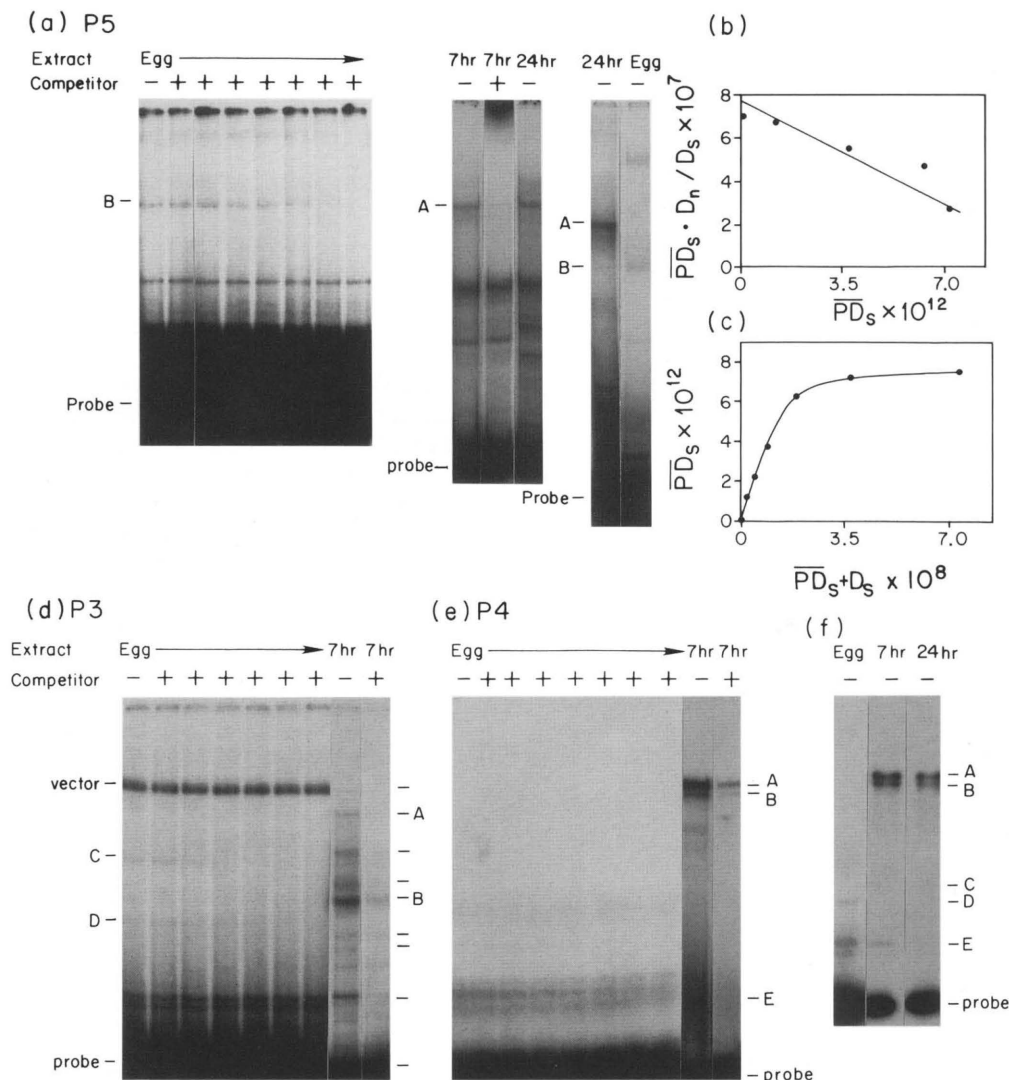


Figure 4. Complexes formed by maternal and embryonic DNA-binding factors P3, P4, and P5. The complete gel-shift competition series are shown for the three sets of factors that were detectable in the egg cytoplasm extracts. Only the complexes and a single illustrative competition lane are shown for qualitative comparison of the complex band pattern in the 7- and 24-hr extract samples. (a) Maternal 7- and 24-hr complexes of factor P5, as indicated. (+) Reactions containing unlabeled competitor. In qualitative comparisons + reactions included 100- to 200-fold molar excess of specific competitor. (b) Least-squares estimate of K_d and P_0 , according to Eq. (1), for maternal complex P5(B). Values from this and a second experiment are shown in Table 2. (c) Saturation curve for maternal complex P5(B), according to Eq. (2). (d) Maternal and 7-hr complexes of factor P3. (e) Maternal and 7-hr complexes of factor P4. (f) Maternal, 7- and 24-hr complexes of P4. This panel was included to show complexes of P4 (C, D) that were not readily visible in e.

ments, the intensity of the P5 interactions is sufficient to permit an estimate of P_o' and K_r . It can be seen in Figure 4a that the maternal complex P5(B) is wholly absent from the 7- and 24-hr gel-shift patterns, in which it is replaced by the P5(A) complex analyzed above (Fig. 2; Table 1). The analysis shown in Figure 4, b and c, demonstrates that the specific affinity (K_r) of the protein forming the P5(B) complex for its site in probe E is the same as for the reaction of the P5(A) protein extracted from the 7- and 24-hr embryo nuclei with probe E. These data are listed in Table 2. Although we do not yet know whether the sequences occupied by the maternal and nuclear P5 factors are wholly identical, an exonuclease III protection experiment indicates that the maternal P5(B) complex begins at the same position (± 5 –10 nucleotides) as the nuclear P5(A) complex (Calzone et al. 1987; additional data not shown). The maternal P_o' value shown for P5(B) is also almost exactly the same as the 24 hr P_o' value for P5(A) (Table 2), i.e., $\sim 10^5$ molecules of each factor per egg or embryo. Thus, it is possible that the maternal P5(B) factor is converted by modification, or removal of a subunit, to the embryonic P5(A) factor. Relatively little P5 protein was recovered from the 7-hr nuclear extract, and what was found is of the P5(A) form. We could not demonstrate either form in 7-hr embryo cytoplasm, but this negative result may be of little significance, as the 7-hr embryo cytoplasmic extracts produced a number of additional nonspecific bands in gel-shift experiments (not shown), which would have obscured detection of specific complexes generated by residual maternal P5 factors. However, neither can we exclude the alternative that the maternal P5 factors are destroyed after fertilization and replaced by new synthesis of the embryonic form, mainly after 7-hr postfertilization (cf. Table 2).

The factors giving rise to the low-prevalence maternal complexes designated P3(C) and P3(D) have disappeared

by 7-hr, whereas the maternal P4(C) complex can still be formed by the 7-hr embryo nuclear extract, although this factor has also disappeared by 24 hr (Fig. 4; Table 2). An important point indicated in Table 2 is that both the P3 and P4 factors have almost achieved the level of prevalence that they display in the 24-hr embryo by 7 hr, i.e., significantly in advance of the activation of the *CyIIIa* gene. Thus, irrespective of the relation between the trace maternal P3 and P4 variants and the P3 and P4 nuclear proteins observed in the nuclear extracts, almost all of the molecules of these factors present in the embryo must be synthesized or released from inaccessible or insoluble structures after fertilization and prior to 7 hr. If their source is de novo synthesis, this could be programmed by maternal or zygotic mRNA, or perhaps both. The gel-shift pattern formed by the 7-hr nuclear extract differs slightly from the 24-hr pattern in that there is an additional specific band that can be competed in the 7-hr extract not seen at 24 hr, namely P4(C), and the ratio of P3(A) to P3(B) is lower at 7 hr. This is unusual, as considered below, for the 7-hr and 24-hr patterns are generally not distinguishable. One possibility is that these changes result from a switch in the 7- to 24-hr interval from maternal to zygotic mRNA templates for the P3 proteins.

Developmental changes in CyIIIa DNA-binding factors that first appear after fertilization

Factors P1, P2, and P6–P8 cannot be detected in unfertilized egg extracts (not shown), but all are present in 7-hr embryo nuclear extracts. Reduced data from gel-shift competitions carried out on the 7-hr embryo extracts are listed in Table 3, where the 7-hr embryo values of K_r and P_o' obtained for these nonmaternal factors are compared with the respective values for the 24-hr embryo extracts from Table 1. In two cases, the 7-hr pat-

Table 2. Prevalence during early development of maternal *CyIIIa*-binding factors

Factor ^a	Unfertilized egg ^b		7 hr ^a		24 hr ^a	
	P_o' ^c	K_r	P_o' ^c	K_r	P_o' ^c	K_r ^d
P3(A)	ND		1.9×10^5	3.6×10^4	3.6×10^5	3.7×10^4
P3(B)	ND		8.3×10^5	3.2×10^4	7.5×10^5	1.7×10^4
P3(C)	+		ND		ND	
P3(D)	+		ND		ND	
P4(A,B)	ND		2.8×10^4	2.8×10^5	2.2×10^5	2.7×10^5
P4(C)	+		+		ND	
P5(A)	ND		5.9×10^3	8.1×10^4	1.7×10^5	7.4×10^4
	ND		5.2×10^3	7.3×10^4		
P5(B)	1.5×10^5	6.8×10^4	ND		ND	
	8.2×10^4	9.4×10^4	ND		ND	

^a The notation for each complex is as shown in Fig. 4.

^b Maternal factors were purified from unfertilized egg proteins soluble at 100,000g, using heparin–agarose columns (see Materials and methods). Factors of 7 and 24 hr were purified from isolated nuclei.

^c Measurements of P_o and K_r were performed as described in Table 1 and in Materials and methods. ND indicates that no complex was detected in the extract. (+) A complex was detected in amounts that were too low to quantify accurately. The minimum amount of P3 and P5 necessary for an accurate determination in these experiments was $\geq 2 \times 10^3$ molecules per embryo. The minimum level necessary for a P4 measurement was approximately 10-fold lower.

^d P_o and K_r values are averages for 24-hr embryo data listed in Table 1.

Table 3. Prevalence at 7 and 24 hr of *CyIIIa*-binding factors that appear after fertilization

Factor	Unfertilized egg	7 hr		24 hr	
	P_o'	P_o'	K_r	P_o'	K_r
P1(A-D)	ND	4.5×10^4	2.4×10^5	4.5×10^5	2.4×10^5
P1(E)	ND	2.3×10^3	5.8×10^5	ND	
P2(A-E)	ND	7.8×10^5	5.4×10^4	2.3×10^6	1.0×10^5
P6	ND	4.6×10^4	4.9×10^5	1.4×10^6	1.2×10^6
P7I(A)	ND	3.4×10^3	5.2×10^5	6.2×10^4	3.3×10^5
P7I(B,C)	ND	1.5×10^3	6.4×10^5	2.5×10^4	1.6×10^6
P7II(A-D)	ND	4.2×10^4	1.0×10^5	7.4×10^5	4.7×10^4
P7II(E)	ND	3.3×10^3	1.0×10^5	ND	
P8(A-C)	ND	4.4×10^5	1.4×10^6	6.1×10^5	9.9×10^5

Probes, competitors, and limits of detection are as in Table 1. The values listed for 24 hr are the averages of the measurements listed in Table 1 for each factor. (ND) Not detected.

terns display an additional complex, i.e., P1(E) and P7II(E). Although these transitory forms are sufficient in concentration for measurement, they are of relatively low prevalence and, thus, resemble the 7-hr P4(C) complex, which is also absent in reactions carried out with the 24-hr extracts as shown in Figure 4e. In all other cases considered in Table 3, i.e., the remaining P1 and P7 complexes, and all the P2, P6, and P8 complexes, as in those considered in Table 2, the patterns of reproducible interactions that can be competed observed with the 7-hr extracts are identical to those observed with the 24-hr nuclear extracts. Furthermore, Table 3 shows that the K_r values obtained for each of these factors in the 7-hr nuclear extracts are indistinguishable from those measured in the 24-hr extracts. Thus, in qualitative terms, this set of *CyIIIa*-binding factors is also established in the embryo after fertilization but prior to activation of the gene. However, it is obviously necessary to regard these only as potential sources of interactions on the *CyIIIa* regulatory region, because these factors are detected in vitro, at various ratios of protein to specific site, and we do not yet know when and where such interactions occur in vivo.

In quantitative terms, an interesting and possibly functional distinction clearly divides these factors into two classes. Table 3 shows that P2(A-E) and P8(A-C) are only two- to threefold more prevalent in the nuclear compartments of the 24-hr embryo than in the nuclear compartments of the 7-hr embryo. This is also true of P3(A,B) and P4(A,B), as shown in Table 2. Because the number of cells per embryo (and of aboral ectoderm precursors) increases about two- to fourfold over this period, the concentrations of these four factors may be the same per nucleus at 7 hr as at 24 hr. In contrast, factors P1(A-D), P6, P7I(A-C), and P7II(A-D) all increase 10- to 40-fold per embryo between 7 and 24 hr. This is also true of the zygotic P5(A) form, but as noted above, P5 could represent a unique case because it is possible that P5(A) derives wholly from the maternal P5(B) form. Because the *CyIIIa* gene begins to be expressed 2-3 hr after the 7-hr time point, this result implies that the rise in concentrations of some or all of these factors to an effec-

tive intranuclear level is the direct cause of temporal activation of this gene.

The various patterns of developmental change in the presence of extractable factors that bind with high specificity to the *CyIIIa* regulators are summarized diagrammatically in Figure 5. Four different patterns may be recognized: (1) *Maternal pattern*: Factor P5(B) is present in eggs at about the same concentration as factor P5(A) in 24-hr embryo nuclei; binding sites for these factors overlap either wholly or partially. Trace quantities of P3(C,D) and P4(C), which disappear early in development, are also detectable in the unfertilized egg. Complexes formed with the maternal factors always migrate differently in gel-shift assays from complexes formed with 24-hr embryo factors. (2) *Early embryonic synthesis pattern*: Several factors, namely P2(A-F), P3(A,B), P4(A,B), and P8(A-C), are not detectable in the egg but have already attained the average levels per nucleus seen at 24 hr by the 7-hr cleavage stage before the *CyIIIa* gene is activated. (3) *Late cleavage/early blastula synthesis pattern*: Factors P1(A-D), P6, P7I(A-C), and P7II(A-D) are all absent from the unfertilized egg and detectable at 7 hr but at only a few percent of the levels observed in 24-hr nuclei. This change appears to be wholly a quantitative one, as the same sets of gel-shift complexes are generated by the extracts from 7- and 24-hr embryo nuclei [if P5(A) is the product of de novo translation rather than a modification of P5(B), it belongs in this class as well]. (4) *Embryonic disappearance of factors*: Three examples of factors present at 7-hr, but absent from 24-hr, nuclear extracts are reported, namely P4(C), P1(E), and P7II(E). The first of these is at a level so low as not to be quantifiable, and the others are present in only minimally detectable quantities, which may not be significant.

It is interesting to consider the patterns that we did not encounter. We saw no examples of maternal factors that generate identical gel-shift bands in extracts from nuclei expressing the *CyIIIa* gene as in extracts from unfertilized eggs. We also saw no qualitatively new bands that appear at 24 hr but are absent from the 7-hr nuclear extracts.

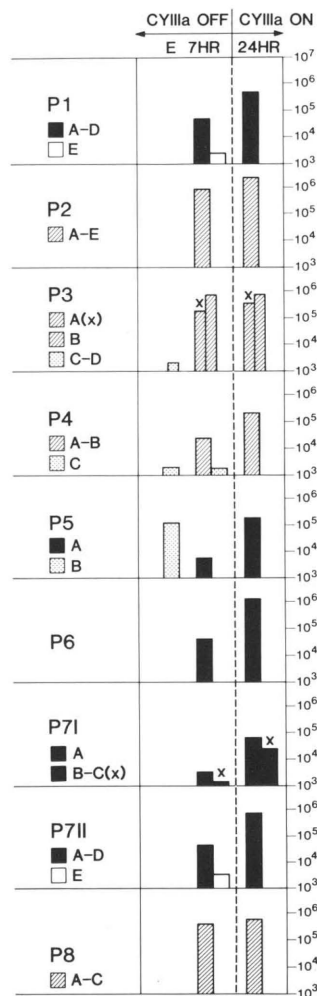


Figure 5. Summary of developmental changes in the prevalence of CyIIIa DNA-binding factors; data are from Tables 1–3. (Top): (E) Maternal factors extracted from unfertilized eggs; (7 hr) factors from 7-hr late cleavage embryo nuclear extracts; (24 hr) factors from 24-hr mesenchyme blastula-stage embryo nuclear extracts. Factor prevalence is calibrated on the right-hand ordinate, on log scales, as molecules per embryo. Solid bars represent factors, the prevalence of which increases by $\geq 10\times$ between 7 and 24 hr, and which are not detectable in unfertilized egg cytoplasm. Open bars represent factors present in 7-hr embryo nuclear extracts only. Hatched bars represent factors not detectable in unfertilized egg cytoplasm, but that, by 7 hr, have already attained a prevalence per embryo in the nuclear compartments $\geq 33\%$ of that observed in 24-hr embryo nuclei. Dotted bars represent maternal factors detected in egg cytoplasm. The sets of gel-shift complexes, or the individual gel-shift complexes utilized for the measurements and calculations in text are considered the products of the individual factors represented to the left of the histogram.

Discussion

Specificity of the CyIIIa DNA–protein interactions

In advance of these measurements, it would have seemed an unlikely prediction that as many as 14 different sequence-specific interactions (Fig. 1) would occur within the regulatory region of a cytoskeletal actin gene expressed in the early embryo. The resolution of the

tests for functionality so far carried out by deletion and in vivo competition (see introductory section) is too low to provide evidence that all, or even a majority, of these individual interactions are either necessary or in any way significant for CyIIIa expression. On the other hand, the gene transfer studies that we have reported certainly demonstrate that the upstream regulatory region as a whole includes those *cis*-interaction sites required for both spatial and temporal regulation of the CyIIIa gene in early development, and the DNA–protein interactions we report here are the only ones that form in the presence of large excess of nonspecific competitor DNA throughout this whole region. It is possible that certain necessary factors have been lost from our extracts, due to unusual lability, or that the lower specificity interactions that cannot be competed, which we ignored in this study, are also important. Barring these caveats, however, it is reasonable to assume that among the set of interactions described in this paper are those that direct the expression of this gene in the embryo.

An issue directly relevant to the significance of these interactions is their sequence specificity. This can be viewed in several ways. Qualitatively, they are sequence specific in that each interaction occurs at a specific, short nucleotide sequence at a particular location in the CyIIIa regulatory DNA. A more convincing and quantitative measure of their specificity arises from the K_d values listed in Table 1. K_d is the key parameter of biological interest in considering putative regulatory functions mediated by DNA–protein interactions, in that this parameter controls the extent of site occupation as a function of the concentration of factor molecules and nonspecific DNA sites available assuming that equilibrium considerations are relevant (see below). The measurements we report here demonstrate preferences ranging from $\sim 3 \times 10^4$ to $\sim 2 \times 10^6$ for the CyIIIa target sequences, with respect to the nonspecific competitors, at least in the milieu of the crude nuclear extracts that we utilized. Given factors may display particular proclivities for competitors of given base composition; e.g., Lin and Riggs (1972, 1975) measured the equilibrium dissociation constant for *lac* repressor interaction with total *Escherichia coli* DNA competitor as $\sim 3 \times 10^{-5}$ M (0.1 M ionic strength), but $\sim 1\text{--}2.6 \times 10^{-6}$ M for poly[d(AT)]. An implication is that the K_d values listed in Table 1 could be minimal for some examples. In any case, the K_d values we measured fall within the range reported for other DNA–binding proteins. Data of Lin and Riggs (1972, 1975) show that against *E. coli* DNA lacking the operator sequence K_d (at 0.1 M ionic strength) is $\sim 4 \times 10^7$, and against poly[d(AT)] is 3×10^6 to 7×10^6 . However, using short probe fragments and gel-shift reactions for the binding assays comparable to those in this work, the *lac* repressor–operator complex displays an equilibrium constant ~ 10 -fold lower than that observed in the earlier measurements (Fried and Crothers 1981). The calculated K_d values are reduced accordingly, and thus, e.g., K_d for *lac* repressor vs. poly[d(AT)] in these determinations is about the same as the average of the K_d values of Table 1 (also measured at

0.1 M K⁺). For comparisons that may be more germane, Emerson et al. (1985) estimated that K_r for a factor that binds specifically to a regulatory region of the chicken β^A -globin gene is $\sim 3.5 \times 10^4$, vs. total *E. coli* DNA; data of Boulanger et al. (1987) indicate K_r is 3.4×10^6 for binding of factor TFIIC2 to its target sequence in the VA1 adenovirus locus, vs. poly[d(I)-d(C)]; and data of Baker et al. (1986) indicate K_r is 1.7×10^5 for the reaction of TFIIA with yeast tRNA DNA, vs. pBR322 DNA (K_{Dn} for the latter two cases is $\sim 1.6 \times 10^{-5}$ M). Similarly, K_r for the *c-fos* enhancer binding protein is $\sim 5 \times 10^4$ and 1×10^5 , vs. DNA and poly[d(AT)], respectively (Prywes and Roeder 1987). We may conclude that the K_r values we report are fairly typical of well-characterized DNA-protein interactions (although it is certainly possible that secondary interactions with other components in the crude extracts have contributed significantly to the results). Although demonstration of the functional importance of most of the interactions measured awaits further gene transfer experiments, it would therefore seem dangerous to regard any of these interactions a priori as accidental or physiologically meaningless. The *CyIIIa* regulatory system is probably combinatorial, perhaps in the sense that the binding factors interact with one another. Thus, for example, Flytzanis et al. (1987) showed that deletion of the restriction fragment that includes the P5 interaction depresses activity of the *CyIIIa*·CAT construct by 85%. Similarly, R. Franks (unpubl.) found that injection of molar excesses of this restriction fragment or of fragments that include only the P2, P3, P4, or P7I sites, respectively, together with *CyIIIa*·CAT, depresses the activity of the latter severely by competition in vivo. It follows that the gene requires, in combination, at least one of the interactions in each of these regions merely to continue transcription in the 24-hr embryo.

The stability determinations shown in Figure 3 provide a real-time measure of the persistence of the DNA-protein complexes, although in an obviously simplified, artificial milieu. Of course, conditions within the nucleus in the immediate vicinity of the *CyIIIa*-binding sites might either stabilize or destabilize such complexes relative to the arbitrary in vitro conditions we applied. In addition, direct interactions with other regulatory proteins may cause dramatic changes in the k_d for DNA-protein complexes as shown in vitro, e.g., for a factor that binds the ovalbumin gene (Tsai et al. 1987). Nonetheless, it is interesting to consider the observed dynamics of complex dissociation in reference to the tempo of the developmental changes under study. With respect to *CyIIIa* gene initiation, the relevant time scale of these biological changes is hours and minutes. Thus, e.g., initiation on the *CyIIIa* gene when it is being transcribed at its maximum rate during the initial accumulation of the transcript (Lee et al. 1986) must occur every few minutes; and the gene is activated at ~ 9 hr of development, after the seventh or eighth cleavage cycle (see introductory section for references). Figure 3 shows that the dissociation rate for one of the most stable of the complexes we observed is $\sim 2 \times 10^{-4}$ /sec [$t_{1/2}$ 47.9

min; k_d is expected to be very insensitive to temperature over the range from 0°C to 15°C, the temperatures at which these embryos are cultured, (assuming behavior similar to that of *lac* repressor-operator complexes; Riggs et al. 1970)]. Measurements presented in Results (Fig. 3 and text) imply that the difference in K_r for P4 and P6 stems mainly from the different stabilities of the complexes formed in these interactions. Assuming this to be general, the K_r values listed in Table 1 imply that the half-lives of the various specific *CyIIIa* complexes would range from a few minutes to an hour or so. Once activated in the embryo, however, the state of *CyIIIa* expression persists for the whole life of the cells of the aboral ectoderm lineages, i.e., at least several days, and the gene may remain functional for the several weeks of larval growth, because *CyIIIa* transcript concentration remains high in the larva (R. Cameron, unpubl.). Functional transcription complexes must therefore exist in the aboral ectoderm cells for periods that are hundreds or thousands times greater than the estimated half-lives of the DNA-protein complexes. Thus, if they are required for activity, these complexes are either enormously more stable in vivo than in vitro, or they are in a continuous state of dynamic equilibrium. The latter seems the more likely possibility, because throughout later development the concentrations of the *CyIIIa* factors in the embryo nuclei remain about the same as in the 24-hr embryo nuclei (N. Thézé, unpubl.).

Assuming equilibrium parameters, would the CyIIIa factors be bound within the embryo nuclei?

Because we do not know to what extent the DNA-binding factors might be compartmentalized within the nuclei, how the stability of their complexes differs from that measured in vitro, or what fraction of genomic DNA is available for nonspecific interaction, the following is largely a heuristic calculation. However, it is useful for interpretation of the P_o' values we observed for the *CyIIIa* factors and also as another illustration of a principle put forth years ago by Lin and Riggs (1975): Because specific complex stabilities are unlikely to be much greater for eukaryotic than for prokaryotic regulatory proteins, in order to achieve specific regulatory interactions in eukaryotic cells, a large number of molecules of regulatory proteins of each species per nucleus might be necessary, compared with the number in prokaryote cells (not considering any exclusion of genomic DNA due to engagement in nucleosomal structures). As we have seen, the K_r values observed for the *CyIIIa* factors are, in fact, not very different from those measured for the *lac* repressor, which is present at ~ 10 molecules per cell. We now assume that there are necessary *CyIIIa* factors confined to the ~ 200 aboral ectoderm precursor cells of the 24-hr embryo; that the volume of each nucleus is $4 \mu\text{m}^3$, and the aggregate volume of the aboral ectoderm nuclei is thus 0.8×10^{-12} liters; that for a typical necessary factor the K_r is 10^5 , and P_o' is 5×10^5 molecules distributed among the 200 nuclei. For a small number of specific sites per haploid genome, e.g., ≤ 10

(i.e., the number of genes in an aboral ectoderm battery utilizing the same regulatory factors) $D_s \cdot K_r$ is insignificant compared with any reasonable estimate of D_n , the concentration of genomic DNA available for nonspecific interaction (usual a priori estimates are 1–10% of the total DNA; e.g., see Emerson et al. 1985; Baker et al. 1986). Thus, Eq. (2) reduces to $PD_s/D_s = P_0 \cdot K_r/D_n$, irrespective of the exact number of sites; and if we set the condition that 90% of these sites be bound at equilibrium, the above values would indicate that 15% of the genomic DNA would serve as nonspecific competitor. This is *prima facie*, perhaps not unreasonable, amounting to about half the internucleosomal DNA. In any case, it can be seen that the measured values of K_r and P_0' are of the appropriate order of magnitude, so that if the in vitro equilibrium conditions were to prevail in vivo, the regulatory sites would indeed be largely occupied by their factors. However, it follows that factors present at much lower levels than 10^5 per embryo early in development, e.g., P7(A–C) or P5(A), which are both recovered at $\leq 10^4$ molecules per embryo in the 7-hr embryo nuclear extract (Fig. 5), are likely to be insufficient in concentration to promote binding, unless they are sequestered in the same few percent of the nuclear volume where the CyIIIa genes reside.

Temporal regulation

The activation of the CyIIIa gene in the late cleavage-stage embryo is a positive regulatory event. Thus, the gene is transcriptionally silent until it is productively expressed (Shott et al. 1984; Lee et al. 1986; Hickey et al. 1987), and the deletions tested by gene transfer (Flytzanis et al. 1987), as well as in vivo competitions (Livant et al. 1988; R. Franks, unpubl.), all depress rather than stimulate CyIIIa · CAT expression. The large increases in concentration per embryo in some or all of the factors responsible for the P1, P6, P7I, and P7II complexes (Fig. 5, solid bars) occurring after 7 hr, i.e., about 2 hr prior to the onset of transcriptional activation, could provide a causal explanation for temporal activation. Accumulation of the number of factor molecules indicated in our measurements would require only a relatively short time even if the mRNAs coding for them were of very low prevalence. Thus, suppose that in the 7 hr following fertilization, 10^6 factor molecules (cf. Fig. 5) were required to be generated, so as to provide $\sim 10^4$ per nucleus in the ~ 100 cell embryo; at the average rate of translation for *S. purpuratus*, 1.8 codons/sec (Goustin and Wilt 1981; for review, see Davidson 1986, pp. 75–78), this could be accomplished by about 1100 active mRNAs in 7 hr (we assume 150 nucleotide spacing between ribosomes, or ~ 27.7 sec/protein molecule released per polysome). Thus, only ~ 10 functional mRNAs coding for this factor would be present in the cytoplasmic volume inherited by a cell at the sixth to seventh cleavage, i.e., probably only $\sim 2 \times 10^{-5}$ of its total mRNA. mRNAs coding for factors that are synthesized mainly after 7 hr are likely to be of zygotic origin, as most maternal mRNAs are being fully utilized by the four- to eight-cell

stage (for review, see Davidson 1986, pp. 77–88). An interesting generalization implied is that the embryo genome may be directly responsible for turning on the relatively large set of specific transcriptional activities that appear at the early blastula stage by expressing a group of regulatory genes required for their temporal activation during cleavage. Such regulatory factors may service genes of many different kinds and could be synthesized in many or all of the embryonic cells.

Materials and methods

DNA probes

The 5'-flanking genomic fragments of the regulatory region of the *S. purpuratus* actin gene CyIIIa (Akhurst et al. 1987) employed as probes for the DNA-binding reactions described are mapped in Figure 1. Each fragment was cloned into pUC18/19 (Stratagene). The probes were labeled by the end-fill reaction with Klenow fragment *E. coli* DNA polymerase, using all four 32 P-labeled deoxynucleotide triphosphates (3000 Ci/mmol, Dupont/NEN), and purified by gel electrophoresis after digestion with restriction enzymes, as indicated in Figure 1.

Isolation of nuclei

S. purpuratus embryos were cultured by standard methods. The fertilization membrane was removed from embryos harvested at 7 hr of development using 1 mM 3-amino-1,2,4-triazole (Sigma) (Showman and Foerder 1979). Embryos (usually 6×10^6) were collected by filtration with 51- μ m Nitex filters and washed $1-2 \times$ by low-speed centrifugation with ice-cold Ca- and Mg-free sea water containing 10 mM Tris (pH 7.4) and 1 mM EDTA. The embryo pellet was resuspended in 10–20 times the pellet volume of TEESD [10 mM Tris (pH 7.4), 1 mM EGTA, 1 mM EDTA, 1 mM spermidine-Tri-HCl, 1 mM dithiothreitol (DTT), 0.36 M sucrose], frozen in liquid N_2 , and stored at 70°C for as long as a year before isolation of nuclei. Cells were lysed by vigorous shaking during thawing. Nuclei were washed $2-3 \times$ in TEESD, and $2 \times$ in TEESD to which 0.1% Triton X-100 was added. The nuclei were transferred to 35-ml 'Oak Ridge' tubes and extracted as described in Parker and Topol (1984). After centrifugation at 3000g, the nuclei were resuspended in 5–10 \times the pellet volume of HEESD [10 mM HEPES (pH 7.9), 1 mM EDTA, 1 mM EGTA, 1 mM spermidine-Tri-HCl, 1 mM DTT]. While mixing the nuclear suspension, one-tenth volume of 4 M ammonium sulfate (pH 7.9) was added dropwise (final concentration 0.36 M). After incubation for 30–60 min on ice, chromatin was removed by centrifugation in a Beckman 60Ti rotor at 35,000 rpm for 1–1.5 hr at 4°C. The chromatin pellets obtained after protein extraction were stored at -20°C for later DNA determinations. Protein was precipitated from the supernatant by the addition of 0.3 g/ml ammonium sulfate. After incubation on ice overnight, the protein precipitate was collected by centrifugation at 10,000g, and dissolved in 0.5 \times the nuclear pellet volume of buffer C [20 mM HEPES (pH 7.9), 40 mM KCl, 0.1 mM EDTA, 1.0 mM DTT, 20% glycerol]. The proteins were dialyzed against buffer C overnight at 4°C. Insoluble proteins were removed by centrifugation, and the extracts were stored in small volumes at -70°C .

For DNA determinations, the chromatin pellets were solubilized by Dounce homogenization in 0.1 M NaCl, 50 mM Tris (pH 7.4), 1 mM EDTA, 1.0% SDS, and 1 mg/ml proteinase K. After incubation overnight at 37°C, DNA concentrations were determined by the DAPI method (Brunk et al. 1979).

Extraction of egg proteins

We devised a method for extraction of egg cytoplasm DNA-binding proteins that consists essentially of concentration of any such proteins from the centrifuged supernatant of a homogenate, by heparin-agarose chromatography. The removal of yolk was necessary prior to the heparin-agarose chromatography step, and any factors bound to the yolk would have been lost from the fractionation. Unfertilized eggs (1×10^7) were washed once with Ca- and Mg-free sea water containing 10 mM Tris (pH 7.4) and 1 mM EDTA. Eggs were resuspended in $10 \times$ the pellet volume of ice-cold buffer C and passed through an 18-gauge needle until completely lysed. The homogenate was centrifuged at 10,000g for 20 min, and the resulting supernatant was centrifuged for 10 min at 35,000 rpm in a Beckman 60Ti rotor to remove additional yolk. After addition of Triton X-100 to 0.5%, ribosomes were removed from the soluble egg proteins by centrifugation at 35,000 rpm in a Beckman 60Ti rotor for 1.5 hr at 4°C. The egg proteins were loaded on a 2-ml heparin-agarose (BRL) column equilibrated in buffer C. After washing the column extensively with buffer C, proteins were eluted in two steps at 0.3 and 0.6 M KCl. Approximately 90% of the eluted protein was present in the 0.3 M KCl fraction. Control experiments carried out with 24-hr embryo nuclear extracts, prepared as above and for which quantitative measurements of factor prevalence had been made, showed that the recovery efficiency of the heparin-agarose elution step was 50-80%. No factors that react with the *CyIIIa* gene were detected in the 0.6 M eluate fraction. Proteins in the 0.3 M KCl fraction were frozen in small aliquots at -70°C .

Gel retardation assays

Each 10- μl binding reaction contained 20 mM HEPES (pH 7.9), 0.5 mM DTT, 75 mM KCl, 5 mM MgCl_2 , a one-fifth dilution of buffer C, 4–6 μg of a synthetic DNA duplex of poly[d(IC)], poly[d(AT)], or poly[dI]/poly[dC] (Pharmacia), 1–4 μg of DNA-binding protein extract, and 0.1–0.5 ng of probe, as indicated in Results. The type and amount of synthetic DNA used as a non-specific competitor, the amount of probe, and the amount of protein added to each reaction were optimized for each factor under analysis. That is, in each case, we tested the three synthetic homopolymers as nonspecific competitors. All behaved equivalently for all the probes, except for the P5 interaction. For this case, we chose poly[dI]/poly[dC], which displayed the lowest value of K_d . The amount of nonspecific DNA included in each reaction, was such as to prevent the formation of significant nonspecific protein-probe complexes, as determined in preliminary trials. An amount of probe was then established that would provide sufficient signal so that the specific competition reactions could be easily quantitated. After incubation at 13°C for 10 min, 1 μl of gel sample buffer (15% Ficoll 400, 0.25% bromphenol blue, 0.25% xylene cyanol) was added to each reaction and the complexes were separated from unbound probe by electrophoresis in 5% acrylamide gels (30% acrylamide/0.8% bis-acrylamide) in $1 \times$ TBE [50 mM Tris-borate (pH 8.3), 1 mM EDTA] at room temperature for 2–4 hr at 200 V, depending on the size of each probe. Gels were prerun at 200 V for 1–2 hr prior to loading each sample. After electrophoresis, the gels were dried, and the radioactive complexes and unbound probe were located by autoradiography. The fraction of probe present in complex was determined by excising the probe or complex and counting each in a scintillation counter. Complexes containing ≤ 10 cpm were quantitated by densitometry of the autoradiograms.

Acknowledgments

We thank Mr. Patrick Leahy for his expert assistance with egg and embryo preparations, and Drs. Norman Davidson and Ellen Rothenberg for helpful and critical reviews of the manuscript. This research was supported by grants from the Lucille P. Markey Charitable Trust and from the National Institutes of Health (HD-05753).

References

- Akhurst, R. J., F. J. Calzone, J. J. Lee, R. J. Britten, and E. H. Davidson. 1987. Structure and organization of the *CyIII* actin gene subfamily of the sea urchin, *Strongylocentrotus purpuratus*. *J. Mol. Biol.* **194**: 193–203.
- Angerer, R. C. and E. H. Davidson. 1984. Molecular indices of cell lineage specification in the sea urchin embryo. *Science* **226**: 1153–1160.
- Baker, R. E., O. Gabrielsen, and B. D. Hall. 1986. Effects of tRNA^{Tyr} point mutations on the binding of yeast RNA polymerase III transcription factor C. *J. Biol. Chem.* **261**: 5275–5282.
- Boulanger, P. A., S. K. Yoshinaga, and A. J. Berk. 1987. DNA-binding properties and characterization of human transcription factor TFIIIC2. *J. Biol. Chem.* **262**: 15098–15105.
- Brunk, C. F., K. C. Jones, and T. W. James. 1979. Assay for nanogram quantities of DNA in cellular homogenates. *Anal. Biochem.* **92**: 497–500.
- Bruskin, A. M., A. Tyner, D. E. Wells, R. M. Showman, and W. H. Klein. 1981. Accumulation in embryogenesis of five mRNAs enriched in the ectoderm of the sea urchin pluteus. *Dev. Biol.* **87**: 308–318.
- Bruskin, A., P. Bedard, A. Tyner, R. Showman, B. Brandhorst, and W. Klein. 1982. A family of proteins accumulating in ectoderm of sea urchin embryos by two related cDNA clones. *Dev. Biol.* **91**: 317–324.
- Calzone, F. J., C. N. Flytzanis, D. R. Fromson, R. J. Britten, and E. H. Davidson. 1987. In *Molecular approaches to developmental biology* (ed. R. A. Firtel and E. H. Davidson), pp. 205–221. Alan R. Liss, New York.
- Cameron, R. A., B. R. Hough-Evans, R. J. Britten, and E. H. Davidson. 1987. Lineage and fate of each blastomere of the eight-cell sea urchin embryo. *Genes Dev.* **1**: 75–85.
- Cox, K. H., L. M. Angerer, J. J. Lee, R. J. Britten, E. H. Davidson, and R. C. Angerer. 1986. Cell lineage-specific programs of expression of multiple actin genes during sea urchin embryogenesis. *J. Mol. Biol.* **188**: 159–172.
- Czihak, G. 1963. Entwicklungsphysiologische Untersuchungen an Echiniden (Verteilung und Bedeutung der Cytochromoxydase). *Roux Archiv. Entwicklungsmech.* **154**: 272–292.
- Davidson, E. H. 1986. *Gene activity in early development*. Academic Press, Orlando, Florida.
- Emerson, B. M., C. D. Lewis, and G. Felsenfeld. 1985. Interaction of specific nuclear factors with the nuclease-hypersensitive region of the chicken adult β -globin gene: Nature of the binding domain. *Cell* **41**: 21–30.
- Flytzanis, C. N., R. J. Britten, and E. H. Davidson. 1987. Ontogenic activation of a fusion gene introduced into the sea urchin egg. *Proc. Natl. Acad. Sci.* **84**: 151–155.
- Franks, R. R., B. R. Hough-Evans, R. J. Britten, and E. H. Davidson. 1988. Spatially deranged though temporally correct expression of a *Strongylocentrotus purpuratus* actin gene fusion in transgenic embryos of a different sea urchin family. *Genes Dev.* **2**: 1–12.
- Fried, M. and D. M. Crothers. 1981. Equilibria and kinetics of

- lac* repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**: 6505–6525.
- Goustin, A. S. and F. H. Wilt. 1981. Protein synthesis, polyribosomes, and peptide elongation in early development of *Strongylocentrotus purpuratus*. *Dev. Biol.* **82**: 32–40.
- Hickey, R. J., M. F. Boshart, and W. R. Crain. 1987. Transcription of three actin genes and a repeated sequence in isolated nuclei of sea urchin embryos. *Dev. Biol.* **124**: 215–227.
- Hough-Evans, B. R., R. J. Britten, and E. H. Davidson. 1988. Mosaic incorporation and regulated expression of an exogenous gene in the sea urchin embryo. *Dev. Biol.* **129** (in press).
- Hough-Evans, B. R., R. R. Franks, R. A. Cameron, R. J. Britten, and E. H. Davidson. 1987. Correct cell type-specific expression of a fusion gene injected into sea urchin eggs. *Dev. Biol.* **121**: 576–579.
- Jobe, A., J. R. Sadler, and S. Bourgeois. 1974. *lac* repressor-operator interaction. IX. The binding of *lac* repressor to operators containing O^c mutations. *J. Mol. Biol.* **85**: 231–248.
- Lee, J. J. 1986. The genomic organization and expression of the *Strongylocentrotus purpuratus* actin gene family. Ph.D. thesis, California Institute of Technology, Pasadena, California.
- Lee, J. J., F. J. Calzone, R. J. Britten, R. C. Angerer, and E. H. Davidson. 1986. Activation of sea urchin actin genes during embryogenesis. Measurement of transcript accumulation from five different genes in *Strongylocentrotus purpuratus*. *J. Mol. Biol.* **188**: 173–183.
- Lin, S.-Y. and A. D. Riggs. 1972. *lac* Repressor binding to non-operator DNA: Detailed studies and a comparison of equilibrium and rate competition methods. *J. Mol. Biol.* **72**: 671–690.
- . 1975. The general affinity of *lac* repressor for *E. coli* DNA: Implications for gene regulation in procaryotes and eucaryotes. *Cell* **4**: 107–111.
- Livant, D., A. Cutting, R. J. Britten, and E. H. Davidson. 1988. Titration of the activity of a fusion gene in intact sea urchin embryos. *Proc. Natl. Acad. Sci.* **85** (in press).
- Parker, C. S. and J. Topol. 1984. A *Drosophila* RNA polymerase II transcription factor binds to the regulatory site of an *hsp70* gene. *Cell* **37**: 273–283.
- Prywes, R. and R. G. Roeder. 1987. Purification of the *c-fos* enhancer-binding protein. *Mol. Cell Biol.* **7**: 3482–3489.
- Riggs, A. D., S. Bourgeois, and M. Cohn. 1970. The *lac* repressor-operator interaction. III. Kinetic studies. *J. Mol. Biol.* **53**: 401–417.
- Shott, R. J., J. J. Lee, R. J. Britten, and E. H. Davidson. 1984. Differential expression of the actin gene family of *Strongylocentrotus purpuratus*. *Dev. Biol.* **101**: 295–306.
- Showman, R. M. and C. A. Foerder. 1979. Removal of the fertilization membrane of sea urchin embryos employing aminotriazole. *Exp. Cell Res.* **120**: 253–255.
- Tsai, S. Y., I. Sagami, H. Wang, M.-J. Tsai, and B. W. O'Malley. 1987. Interactions between a DNA-binding transcription factor (COUP) and a non-DNA binding factor (S300-II). *Cell* **50**: 701–709.



Developmental appearance of factors that bind specifically to cis-regulatory sequences of a gene expressed in the sea urchin embryo.

F J Calzone, N Thézé, P Thiebaud, et al.

Genes Dev. 1988 2: 1074-1088

Access the most recent version at doi:[10.1101/gad.2.9.1074](https://doi.org/10.1101/gad.2.9.1074)

References

This article cites 27 articles, 7 of which can be accessed free at:
<http://genesdev.cshlp.org/content/2/9/1074.full.html#ref-list-1>

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Genes & Development* go to:
<http://genesdev.cshlp.org/subscriptions>
